

PCT09

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/868,195

DATE: 01/15/2002

TIME: 12:56:33

Input Set : A:\GJE-71.ST25.txt

Output Set: N:\CRF3\01152002\I868195.raw

ENTERED

3 <110> APPLICANT: Hughes, Martin J G
 4 Santangelo, Joseph D
 5 Lane, Jonathan D
 6 Feldman, Robert
 7 Moore, Joanne C
 8 Dobson, Richard J
 9 Everest, Paul
 10 Henwood, Caroline J
 11 Dougan, Gordon
 12 Wilson, Rebecca K
 14 <120> TITLE OF INVENTION: Outer Surface Proteins, Their Genes, and Their Use
 16 <130> FILE REFERENCE: GJE-71
 18 <140> CURRENT APPLICATION NUMBER: US 09/868,195
 19 <141> CURRENT FILING DATE: 2001-06-15
 21 <160> NUMBER OF SEQ ID NOS: 12
 23 <170> SOFTWARE: PatentIn version 3.1
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 1014
 27 <212> TYPE: DNA
 28 <213> ORGANISM: Streptococcus agalactiae
 30 <220> FEATURE:
 31 <221> NAME/KEY: CDS
 32 <222> LOCATION: (1)..(1014)
 33 <223> OTHER INFORMATION:
 36 <400> SEQUENCE: 1
 37 atg aca caa gta ttt caa gga cgt agt ttc tta gca gaa aaa gat ttt 48
 38 Met Thr Gln Val Phe Gln Gly Arg Ser Phe Leu Ala Glu Lys Asp Phe
 39 1 5 10 15
 41 tct cgt gag gaa ttt gaa tat ctt att gat ttt tca gct cat tta aaa 96
 42 Ser Arg Glu Glu Phe Glu Tyr Leu Ile Asp Phe Ser Ala His Leu Lys
 43 20 25 30
 45 gac ctt aaa aaa cgt ggt gtt cct cat cat tat ctt gaa ggt aaa aat 144
 46 Asp Leu Lys Lys Arg Gly Val Pro His His Tyr Leu Glu Gly Lys Asn
 47 35 40 45
 49 att gct ctc tta ttt gaa aaa aca tct act cgt act cgc gca gcc ttt 192
 50 Ile Ala Leu Leu Phe Glu Lys Thr Ser Thr Arg Thr Arg Ala Ala Phe
 51 50 55 60
 53 aca act gca gca att gac cta ggc gct cat ccg gaa tac ctt ggt gca 240
 54 Thr Thr Ala Ala Ile Asp Leu Gly Ala His Pro Glu Tyr Leu Gly Ala
 55 65 70 75 80
 57 aat gat att caa ctt ggt aaa aaa gaa tca aca gaa gat act gct aag 288
 58 Asn Asp Ile Gln Leu Gly Lys Lys Glu Ser Thr Glu Asp Thr Ala Lys
 59 85 90 95
 61 gtt tta gga cgt atg ttt gat ggt att gaa ttc cgt ggt ttt agc caa 336
 62 Val Leu Gly Arg Met Phe Asp Gly Ile Glu Phe Arg Gly Phe Ser Gln
 63 100 105 110
 65 aga atg gtt gaa gag ctt gct gaa ttt tct gga gta cct gtc tgg aat 384

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/868,195

DATE: 01/15/2002

TIME: 12:56:33

Input Set : A:\GJE-71.ST25.txt

Output Set: N:\CRF3\01152002\I868195.raw

```

66 Arg Met Val Glu Glu Leu Ala Glu Phe Ser Gly Val Pro Val Trp Asn
67      115      120      125
69 ggt tta aca gat gaa tgg cat cca aca caa atg cta gct gac tac ctt      432
70 Gly Leu Thr Asp Glu Trp His Pro Thr Gln Met Leu Ala Asp Tyr Leu
71      130      135      140
73 act atc aaa gaa aac ttc ggt aaa ctt gaa ggt att act ctt gtt tac      480
74 Thr Ile Lys Glu Asn Phe Gly Lys Leu Glu Gly Ile Thr Leu Val Tyr
75      145      150      155      160
77 tgt ggt gac gga cgt aac aat gtt gcc aac tcg ctt tta gtg gct ggg      528
78 Cys Gly Asp Gly Arg Asn Asn Val Ala Asn Ser Leu Leu Val Ala Gly
79      165      170      175
81 act ttg atg ggg gtc aat gta cac atc ttt tct cca aaa gaa ctt tty      576
82 Thr Leu Met Gly Val Asn Val His Ile Phe Ser Pro Lys Glu Leu Phe
83      180      185      190
85 ccw gct gaa gag att gtt aaa ttg gct gaa gga tat gcc aaa gaa tct      624
86 Pro Ala Glu Glu Ile Val Lys Leu Ala Glu Gly Tyr Ala Lys Glu Ser
87      195      200      205
89 ggg gct cac gtt ctc gtt act gat aat gta gac gaa gct gta aag gga      672
90 Gly Ala His Val Leu Val Thr Asp Asn Val Asp Glu Ala Val Lys Gly
91      210      215      220
93 gca gac gtc ttt tac act gat gtc tgg gta tcg atg gga gaa gaa gat      720
94 Ala Asp Val Phe Tyr Thr Asp Val Trp Val Ser Met Gly Glu Glu Asp
95      225      230      235      240
97 aag ttc aaa gaa cgc gtt gaa ctt ctt caa cca tat caa gta aac atg      768
98 Lys Phe Lys Glu Arg Val Glu Leu Leu Gln Pro Tyr Gln Val Asn Met
99      245      250      255
101 gaa ctg att aaa aaa gct aat aat gat aat ctt atc ttc tta cac tgc      816
102 Glu Leu Ile Lys Lys Ala Asn Asn Asp Asn Leu Ile Phe Leu His Cys
103      260      265      270
105 tta cct gca ttc cat gat aca aat acc gtt tat ggc aaa gac gtc gct      864
106 Leu Pro Ala Phe His Asp Thr Asn Thr Val Tyr Gly Lys Asp Val Ala
107      275      280      285
109 gaa aaa ttt ggg gtc aag gaa atg gaa gtt act gat gaa gtc ttc cgt      912
110 Glu Lys Phe Gly Val Lys Glu Met Glu Val Thr Asp Glu Val Phe Arg
111      290      295      300
113 agc aaa tat gct cgt cat ttc gac caa gct gaa aat cgt atg cac act      960
114 Ser Lys Tyr Ala Arg His Phe Asp Gln Ala Glu Asn Arg Met His Thr
115      305      310      315      320
117 att aaa gct gta atg gct gca acc ctt gga aat ctt ttc att cca aaa      1008
118 Ile Lys Ala Val Met Ala Ala Thr Leu Gly Asn Leu Phe Ile Pro Lys
119      325      330      335
121 gtt taa      1014
122 Val
126 <210> SEQ ID NO: 2
127 <211> LENGTH: 337
128 <212> TYPE: PRT
129 <213> ORGANISM: Streptococcus agalactiae
132 <400> SEQUENCE: 2
134 Met Thr Gln Val Phe Gln Gly Arg Ser Phe Leu Ala Glu Lys Asp Phe

```

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/868,195

TIME: 12:56:33

Input Set : A:\GJE-71.ST25.txt

Output Set: N:\CRF3\01152002\I868195.raw

```

135 1          5          10          15
138 Ser Arg Glu Glu Phe Glu Tyr Leu Ile Asp Phe Ser Ala His Leu Lys
139          20          25          30
142 Asp Leu Lys Lys Arg Gly Val Pro His His Tyr Leu Glu Gly Lys Asn
143          35          40          45
146 Ile Ala Leu Leu Phe Glu Lys Thr Ser Thr Arg Thr Arg Ala Ala Phe
147          50          55          60
150 Thr Thr Ala Ala Ile Asp Leu Gly Ala His Pro Glu Tyr Leu Gly Ala
151 65          70          75          80
154 Asn Asp Ile Gln Leu Gly Lys Lys Glu Ser Thr Glu Asp Thr Ala Lys
155          85          90          95
158 Val Leu Gly Arg Met Phe Asp Gly Ile Glu Phe Arg Gly Phe Ser Gln
159          100          105          110
162 Arg Met Val Glu Glu Leu Ala Glu Phe Ser Gly Val Pro Val Trp Asn
163          115          120          125
166 Gly Leu Thr Asp Glu Trp His Pro Thr Gln Met Leu Ala Asp Tyr Leu
167          130          135          140
170 Thr Ile Lys Glu Asn Phe Gly Lys Leu Glu Gly Ile Thr Leu Val Tyr
171 145          150          155          160
174 Cys Gly Asp Gly Arg Asn Asn Val Ala Asn Ser Leu Leu Val Ala Gly
175          165          170          175
178 Thr Leu Met Gly Val Asn Val His Ile Phe Ser Pro Lys Glu Leu Phe
179          180          185          190
182 Pro Ala Glu Glu Ile Val Lys Leu Ala Glu Gly Tyr Ala Lys Glu Ser
183          195          200          205
186 Gly Ala His Val Leu Val Thr Asp Asn Val Asp Glu Ala Val Lys Gly
187          210          215          220
190 Ala Asp Val Phe Tyr Thr Asp Val Trp Val Ser Met Gly Glu Glu Asp
191 225          230          235          240
194 Lys Phe Lys Glu Arg Val Glu Leu Leu Gln Pro Tyr Gln Val Asn Met
195          245          250          255
198 Glu Leu Ile Lys Lys Ala Asn Asn Asp Asn Leu Ile Phe Leu His Cys
199          260          265          270
202 Leu Pro Ala Phe His Asp Thr Asn Thr Val Tyr Gly Lys Asp Val Ala
203          275          280          285
206 Glu Lys Phe Gly Val Lys Glu Met Glu Val Thr Asp Glu Val Phe Arg
207          290          295          300
210 Ser Lys Tyr Ala Arg His Phe Asp Gln Ala Glu Asn Arg Met His Thr
211 305          310          315          320
214 Ile Lys Ala Val Met Ala Ala Thr Leu Gly Asn Leu Phe Ile Pro Lys
215          325          330          335
218 Val
222 <210> SEQ ID NO: 3
223 <211> LENGTH: 1197
224 <212> TYPE: DNA
225 <213> ORGANISM: Streptococcus agalactiae
227 <220> FEATURE:
228 <221> NAME/KEY: CDS
229 <222> LOCATION: (1)..(1197)

```

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/868,195

TIME: 12:56:33

Input Set : A:\GJE-71.ST25.txt

Output Set: N:\CRF3\01152002\I868195.raw

```

230 <223> OTHER INFORMATION:
233 <400> SEQUENCE: 3
234 atg gct aaa ttg act gtt aaa gac gtt gat ttg aag gta aaa aaa gtc      48
235 Met Ala Lys Leu Thr Val Lys Asp Val Asp Leu Lys Val Lys Lys Val
236 1          5          10          15
238 ctc gtt cgt gtt gac ttt aat gtg cct ttg aaa gac ggc gtt atc act      96
239 Leu Val Arg Val Asp Phe Asn Val Pro Leu Lys Asp Gly Val Ile Thr
240          20          25          30
242 aac gac aac cgt atc act gcg gct ctt cca aca atc aag tat atc atc      144
243 Asn Asp Asn Arg Ile Thr Ala Ala Leu Pro Thr Ile Lys Tyr Ile Ile
244          35          40          45
246 gaa caa ggt ggt cgt gct atc ctc ttc tct cac ctt gga cgt gtt aaa      192
247 Glu Gln Gly Gly Arg Ala Ile Leu Phe Ser His Leu Gly Arg Val Lys
248          50          55          60
250 gaa gaa gct gac aaa gaa gga aaa tca ctt gca ccg gta gct gct gat      240
251 Glu Glu Ala Asp Lys Glu Gly Lys Ser Leu Ala Pro Val Ala Ala Asp
252 65          70          75          80
254 tta gct gct aaa ctt ggt caa gat gtt gta ttc cca ggt gtt act cgt      288
255 Leu Ala Ala Lys Leu Gly Gln Asp Val Val Phe Pro Gly Val Thr Arg
256          85          90          95
258 ggt gca aaa tta gaa gaa gca atc aat gct ttg gaa gat gga caa gtt      336
259 Gly Ala Lys Leu Glu Glu Ala Ile Asn Ala Leu Glu Asp Gly Gln Val
260          100          105          110
262 ctt ttg gtt gaa aac act cgt ttt gaa gat gtt gac ggt aag aaa gaa      384
263 Leu Leu Val Glu Asn Thr Arg Phe Glu Asp Val Asp Gly Lys Lys Glu
264          115          120          125
266 tct aag aat gac gaa gaa ctt ggt aaa tac tgg gct tca ctt gga gat      432
267 Ser Lys Asn Asp Glu Glu Leu Gly Lys Tyr Trp Ala Ser Leu Gly Asp
268          130          135          140
270 gga atc ttc gtt aac gat gca ttt ggt aca gca cac cgt gct cat gca      480
271 Gly Ile Phe Val Asn Asp Ala Phe Gly Thr Ala His Arg Ala His Ala
272 145          150          155          160
274 tca aac gta ggt att tca gca aac gtt gaa aaa gct gta gct ggt ttc      528
275 Ser Asn Val Gly Ile Ser Ala Asn Val Glu Lys Ala Val Ala Gly Phe
276          165          170          175
278 ctt ctt gaa aac gaa att gct tac atc caa gaa gca gtt gaa act cca      576
279 Leu Leu Glu Asn Glu Ile Ala Tyr Ile Gln Glu Ala Val Glu Thr Pro
280          180          185          190
282 gaa cgc cca ttc gta gct att ctt ggt ggc tca aaa gtt tct gat aag      624
283 Glu Arg Pro Phe Val Ala Ile Leu Gly Gly Ser Lys Val Ser Asp Lys
284          195          200          205
286 att ggt gtt atc gaa aac ctt ctt gaa aaa gct gat aaa gtt ctt atc      672
287 Ile Gly Val Ile Glu Asn Leu Leu Glu Lys Ala Asp Lys Val Leu Ile
288          210          215          220
290 ggt ggt ggt atg act tac aca ttc tac aaa gct caa ggt atc gaa atc      720
291 Gly Gly Gly Met Thr Tyr Thr Phe Tyr Lys Ala Gln Gly Ile Glu Ile
292 225          230          235          240
294 ggt aac tca ctt gta gaa gaa gac aaa ttg gat gtt gct aaa gac ctc      768
295 Gly Asn Ser Leu Val Glu Glu Asp Lys Leu Asp Val Ala Lys Asp Leu

```

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/868,195

TIME: 12:56:33

Input Set : A:\GJE-71.ST25.txt

Output Set: N:\CRF3\01152002\I868195.raw

```

296          245          250          255
298 ctt gaa aaa tca aac ggt aaa ttg atc ttg cca gtt gac tca aaa gaa      816
299 Leu Glu Lys Ser Asn Gly Lys Leu Ile Leu Pro Val Asp Ser Lys Glu
300          260          265          270
302 gca aac gca ttt gct ggt tat act gaa gtt cgc gac act gaa ggt gaa      864
303 Ala Asn Ala Phe Ala Gly Tyr Thr Glu Val Arg Asp Thr Glu Gly Glu
304          275          280          285
306 gca gtt tca gaa ggg ttc ctt ggt ctt gac atc ggt cct aaa tca atc      912
307 Ala Val Ser Glu Gly Phe Leu Gly Leu Asp Ile Gly Pro Lys Ser Ile
308          290          295          300
310 gct aaa ttt gat gaa gca ctt act ggt gct aaa aca gtt gta tgg aac      960
311 Ala Lys Phe Asp Glu Ala Leu Thr Gly Ala Lys Thr Val Val Trp Asn
312 305          310          315          320
314 gga cct atg ggt gtc ttt gaa aac cct gac ttc caa gct ggt aca atc      1008
315 Gly Pro Met Gly Val Phe Glu Asn Pro Asp Phe Gln Ala Gly Thr Ile
316          325          330          335
318 ggt gta atg gac gct atc gtt aaa caa cca ggc gtt aaa tca atc atc      1056
319 Gly Val Met Asp Ala Ile Val Lys Gln Pro Gly Val Lys Ser Ile Ile
320          340          345          350
322 ggt ggt ggt gat tca gca gca gct gct atc aac ctt ggt cgt gct gac      1104
323 Gly Gly Gly Asp Ser Ala Ala Ala Ala Ile Asn Leu Gly Arg Ala Asp
324          355          360          365
326 aaa ttc tca tgg atc tct act ggt ggt gga gca agc atg gaa ttg ctc      1152
327 Lys Phe Ser Trp Ile Ser Thr Gly Gly Gly Ala Ser Met Glu Leu Leu
328          370          375          380
330 gaa ggt aaa gta tta cca ggt ttg gca gca ttg act gaa aaa taa      1197
331 Glu Gly Lys Val Leu Pro Gly Leu Ala Ala Leu Thr Glu Lys
332 385          390          395
335 <210> SEQ ID NO: 4
336 <211> LENGTH: 398
337 <212> TYPE: PRT
338 <213> ORGANISM: Streptococcus agalactiae
340 <400> SEQUENCE: 4
342 Met Ala Lys Leu Thr Val Lys Asp Val Asp Leu Lys Val Lys Lys Val
343 1          5          10          15
346 Leu Val Arg Val Asp Phe Asn Val Pro Leu Lys Asp Gly Val Ile Thr
347          20          25          30
350 Asn Asp Asn Arg Ile Thr Ala Ala Leu Pro Thr Ile Lys Tyr Ile Ile
351          35          40          45
354 Glu Gln Gly Gly Arg Ala Ile Leu Phe Ser His Leu Gly Arg Val Lys
355          50          55          60
358 Glu Glu Ala Asp Lys Glu Gly Lys Ser Leu Ala Pro Val Ala Ala Asp
359 65          70          75          80
362 Leu Ala Ala Lys Leu Gly Gln Asp Val Val Phe Pro Gly Val Thr Arg
363          85          90          95
366 Gly Ala Lys Leu Glu Glu Ala Ile Asn Ala Leu Glu Asp Gly Gln Val
367          100          105          110
370 Leu Leu Val Glu Asn Thr Arg Phe Glu Asp Val Asp Gly Lys Lys Glu
371          115          120          125

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/868,195

DATE: 01/15/2002

TIME: 12:56:34

Input Set : A:\GJE-71.ST25.txt

Output Set: N:\CRF3\01152002\I868195.raw